AAPM Joint Imaging/Therapy Symposium
Imaging as a Biomarker

Infrastructure to Integrate Imaging in Clinical Trials

Daniel C. Sullivan, M.D.
Duke University Medical Center
Science Advisor, RSNA
May 31, 2009

RSNA Radiological Society of North America Founded in 1915

Quantitative Imaging Biomarkers Alliance RSNA

Financial Disclosures/COI

- Scientific Advisory Board, MDS Nordion

Evaluation of Imaging Tests

Assumptions:

- Imaging provides information.
- By influencing physician thinking and behavior, this information affects patient outcomes.

Questions:

- How accurate & reliable is the information?
- How valuable is the information?

Clinical Trials of Imaging Tests

A. Specific Aim: Establish Performance Characteristics of Test *x*)

- Observational Study 1
- Observational Study *m*

B. Specific Aim: Correlate Test *x* with Condition *y*)

- Testable Hypothesis 1 (Designed Clinical Trial 1)
- Testable Hypothesis 2 (Designed Clinical Trial 2)
- Testable Hypothesis 3 (Designed Clinical Trial 3)
- Testable Hypothesis *n* (Designed Clinical Trial *n*)

- Cumulative data = "Qualification" (High confidence level in clinical significance of results from Test *x*).

Biomarker Levels of Evidence

I: A. Single, prospective trial in which marker determines clinical decision;

or B. Overview of Level II studies

II: Prospective trials evaluating relationship of marker to proposed utility.

III: Large, retrospective studies.

IV: Small, retrospective studies.

V: Small pilot studies.

Hayes, et.al. 1996

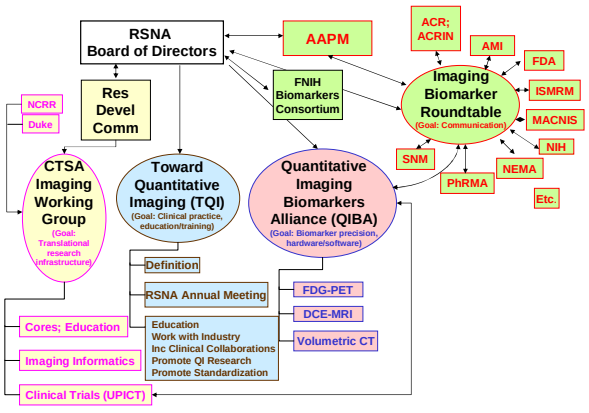
RSNA Interests

- RSNA is interested in fostering more emphasis on quantitative imaging in clinical care
- Facilitating ***imaging as a biomarker in clinical trials*** helps RSNA move this agenda forward.

RSNA Radiological Society of North America Founded in 1915

What are the hurdles to better quantification?

- **Radiologists**
  - Skepticism about clinical value/need
  - Concern about variability (reader & machine)
  - Lack of incentive
- **Scanner Manufacturers**
  - Lack of customer demand
  - Concern about "standardization" (commoditization vs. differentiation)
  - Lack of ROI (Return on Investment)/Opportunity Cost



TQI Quantitative Imaging Reading Room of the Future

- Educational Exhibit Hall, RSNA Annual Meeting 2009
- 15 exhibits related to software that can be incorporated now (or soon) into routine radiologic practice.

CTSA Imaging Working Group

3 Subcommittees:

- Cores (Structure; Administration; Financing)
- Imaging Informatics (Integrate existing tools)
- Clinical Trials (UPICT – Uniform Protocols for Imaging in Clinical Trials)

Imaging Biomarkers Roundtable

- Communication
- Coordinate Activities
- Next Mtg Nov 3-4, 2009

QIBA Background

- Began May, 2008
- Mission: Improve value and practicality of quantitative imaging biomarkers by reducing variability across devices, patients, and time.
  - Build “measuring devices” rather than “imaging devices”.

QIBA Process

- Identify sources of variability
- Collect "groundwork" data
- Devise mitigation strategies
- Write and promulgate "Profiles".

QIBA Progress

- Built on antecedent activities of several other organizations:
  - AAPM
  - IRAT
  - ISMRM
  - SNM
  - FDA
  - NCI
  - ADNI

QIBA Technical Committees

Quantitative Imaging Biomarkers Alliance RSNA

Fluorodeoxyglucose Positron Emission Tomography (FDG-PET/CT)
- Richard Frank, MD, PhD, GE Healthcare
- Helen Young, PhD, AstraZeneca
- Sandy McEwan, MD, Univ of Alberta

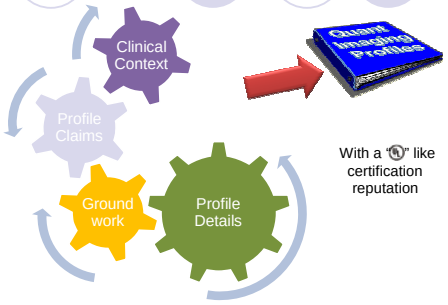
Dynamic Contrast-Enhanced Magnetic Resonance Imaging (DCE-MRI)
- Gudrun Zahlmann, PhD, Roche
- Jeff Evelhoch, PhD, Merck
- Michael Buonocore, PhD, MD, UC Davis

Volumetric Computed Tomography (Vol-CT)
- Andrew Buckler, MS, Buckler Consulting
- David Mozley, MD, Merck
- Larry Schwartz, MD, MSKCC

Factors Affecting QIBA Scope

- NIST definition of a measurement result: "A measurement result is complete only when accompanied by a quantitative statement of its uncertainty. The uncertainty is required in order to decide if the result is adequate for its intended purpose and to ascertain if it is consistent with other similar results."
- FDA: "A biomarker must be qualified for its intended purpose"

The QIBA Process



Result: QIBA Profiles

- **A QIBA Profile is a document with 3 parts.**
- **Part 1 tells a user what can be accomplished by following the Profile. (Claims)**
  - **E.g. "you will be able to detect volume changes of greater than 10% in Stage I lung cancer nodules which are 5mm in diameter or greater."**

QIBA Profile (2)

- **It tells a vendor what they must implement in their product to state compliance with the Profile. (Details)**
  - **E.g. to comply, the scanner must be able to:**
    - **scan a Mark-324 Chest Phantom, identify the smallest resolvable target, display the diameter of that target**
    - **demonstrate resolving targets at least as small as 2mm diameter on the Mark-324 phantom**
    - **scan patients according to the ACRIN NLST acquisition protocol**
  - **E.g. to comply, the quantification application must be able to:**
    - **segment a nodule (automatically or manually), derive the volume, store it in a DICOM object**
    - **run a user through a set of test data with known volumes and at the end display an accuracy score**

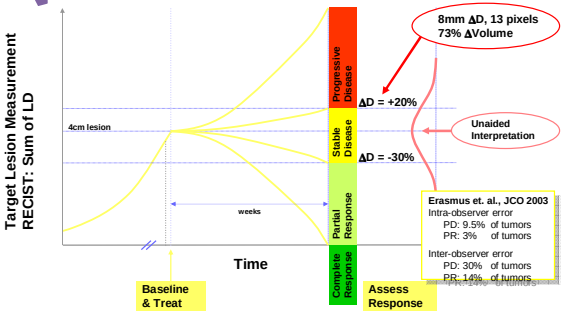
QIBA Profile (3)

- **It tells the user staff what they must do for the Profile Claims to be realized. (Details)**
  - **E.g. to comply, the site CT techs must be able to:**
    - **scan the patient within 10 minutes of contrast injection**
  - **E.g. to comply, the radiologist must be able to:**
    - **achieve a score of 95% or better using their segmentation application on the LIDC test set.**

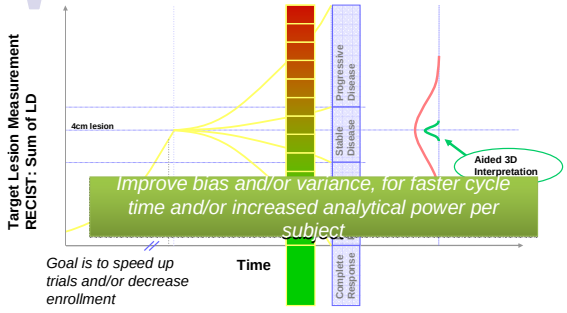
Clinical Context

- Defined by ad-hoc sub-committees of clinicians:
  - Start by determining the clinical context, e.g., disease staging, clinical manifestations, etc.
  - Determine what biomarkers to pursue
  - For each biomarker, determine what Profiles to pursue

Improvement Begins From the Current Standard



Volumetric CT Seeks to Increase Sensitivity vs. LD



Groundwork / Profile Details

- To include explicit coverage of:
  - Covariates rationale, e.g. patient preparation
  - Quality control metrics
  - Acquisition protocols
  - ROI definition
  - Quantification computation
  - Data transfer and storage issues
  - Longitudinal measurements

Ground work

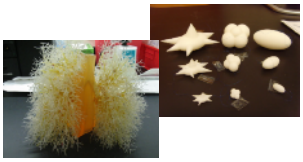
Part 1A Analyzes Bias and Variance where Ground Truth is Known Deterministically

Question: What is reader variation under a ~~limited~~ set of ~~controlled~~ conditions for a reference set of image data?

Approach: Inter/Intra-reader reliability evaluation utilizing a volumetric software tool

Ground truth by physical measurement "ex vivo".

Images acquired by FDA/CDRH/OSEL
•Estimate nodule size from CT images of variety of phantom nodules



Ground work

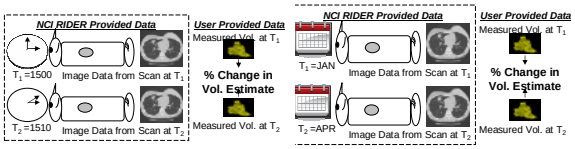
Part 1B Focuses on Change Analysis in Diagnostic Settings

Question: What is measurement variation considered a "no change" condition (which constitutes minimum detectable change)?

Approach: Coffee Break Experiment – provides a close to "no change" condition in vivo.
Publicly Available on NCIA

Question: What is measurement variation for different software/user methods for a reference set of image data?

Approach: Patient Change Data from LIDC Database
Publicly Available on NCIA

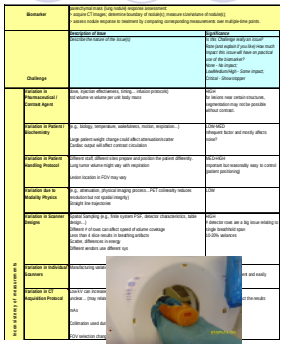


Ground work

Part 1C Measures Multi-center Variability, Sans Biology

Multi-Center Phantom Study to Model Sources of Variability from a Systems Engineering Analysis
Multiple image sets of the same phantoms re-scanned across centers to isolate contributors to variability.

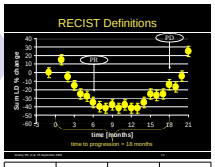
The goal is to determine necessary control conditions to be documented in profiles ensuring that the output for imaging when performed under these conditions will be adequately precise and accurate when scanned on profile-compliant equipment.



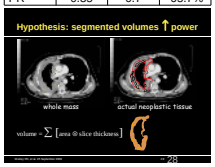
Ground work

Outcome Prediction

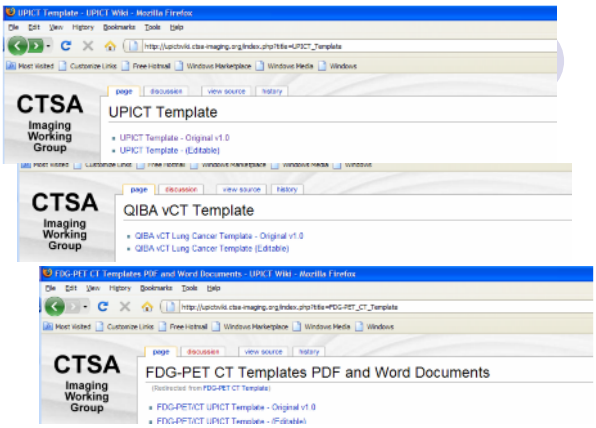
- **Correlation between new biomarker and RECIST**
- **Progress from single to multiple image analysts**
- **Estimate value of new biomarker versus standard in terms of:**
  - **Increased analytical power per subject,**
  - **Length of time each subject needs to stay on trial or a treatment regimen, and**
  - **Cycle time required to make critical GO or NO GO decisions based on group differences between treatment arms in clinical trials.**



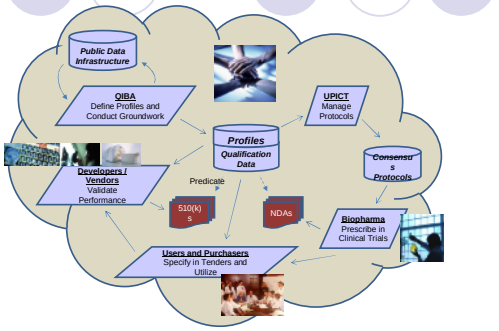
	Radius of Ball	Longest Line-length of Cube	Δ baseline
baseline	0.50	1.0	
PD	0.60	1.2	72.8%
PR	0.35	0.7	-65.7%



Slides shown are illustrative only, from Mozley et al.



We are Working to Integrate the Imaging Biomarker Enterprise



Thank you

Quantitative Imaging Biomarkers Alliance RSNA